

Numerical Evaluation of Diffraction Integrals

Volume 105

Number 4

July–August 2000

Klaus D. Mielenz

Oakland, MD 21550

kmielenz@mail.gcnet.net

This paper describes a simple numerical integration method for diffraction integrals which is based on elementary geometrical considerations of the manner in which different portions of the incident wavefront contribute to the diffracted field. The method is applicable in a wide range of cases as the assumptions regarding the type of integral are minimal, and the results are accurate even when the wavefront is divided into only a relatively small number of summation elements. Higher accuracies can be achieved by increasing the number of summation elements and/or incorporating Simpson's rule into the basic integration formula. The use of the method is illustrated by numerical exam-

ples based on Fresnel's diffraction integrals for circular apertures and apertures bounded by infinite straight lines (slits, half planes). In the latter cases, the numerical integration formula is reduced to a simple recursion formula, so that there is no need to perform repetitive summations for every point of the diffraction profile.

Key words: circular aperture; diffraction; half plane; numerical integration; recursion formula; slit.

Accepted: July 28, 2000

Available online: <http://www.nist.gov/jres>

1. Basic Equations

Diffraction problems that can be described in two dimensions as indicated in Fig. 1 usually lead to complex integrals of the type

$$U(x, z) = \int d\xi A(\xi, x, z) B(\xi, x, z), \quad (1)$$

where A and B account for the amplitude and phase of the optical field at a point of observation $P = (x, z)$ due to a point $Q = (\xi, 0)$ located inside a diffracting aperture. In practice, the phase term B oscillates rapidly inside the range of integration while the amplitude term A varies slowly. The phase term B is often, but not always, a sinusoidal function of the form $e^{ikg(\xi)}$, where $k = 2\pi/\lambda$ is the circular wavenumber of monochromatic light with wavelength λ .

When closed analytical expressions for U are not available it is sometimes possible to find approximate solutions by the method of stationary phase [1,2]. However, in this author's experience, the results obtained can be unreliable. A potential source of error is the basic premise of the method itself; namely, that the rapid oscillations of the phase term B nullify each other except in the vicinity of stationary points. This is presumed true on account of the slow variation of the amplitude term A , but inconsistent with the well-known fact that the corresponding series encountered in connection with Fresnel's zone construction, $A_1 - A_2 + A_3 - \dots - A_n$, has a finite value even when the terms alter their absolute values very gradually [3,4]. Accordingly, substantial errors can occur if the contribution of

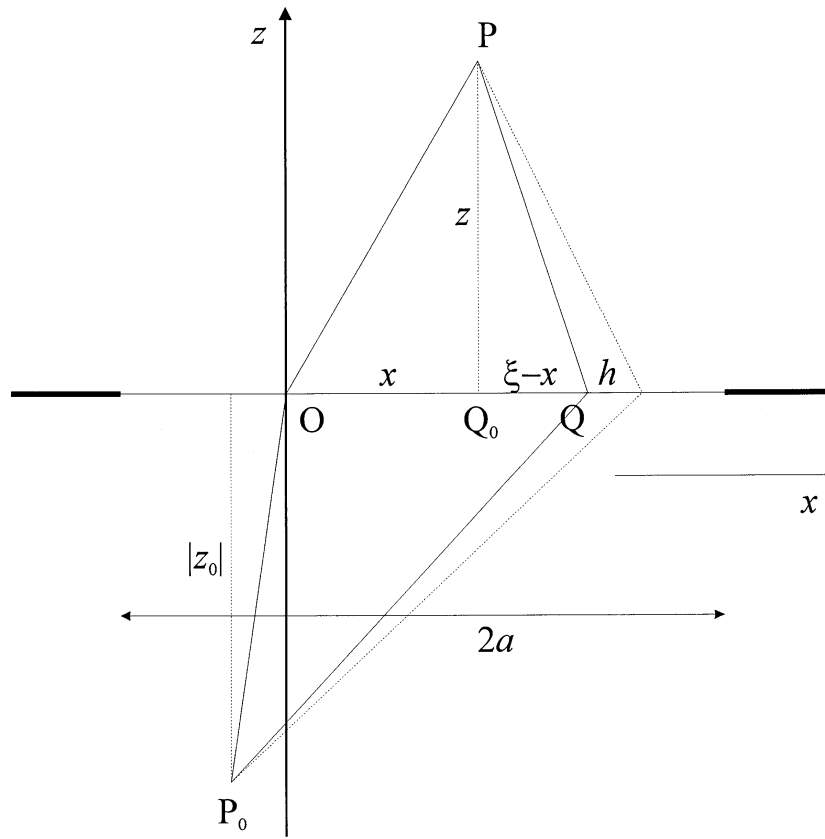


Fig. 1. Cross section of a plane diffracting aperture (xz -plane).

the stationary points is weak. As shown in Appendix A, the stationary-phase method fails completely when applied to Fresnel diffraction at a slit or half plane.

For these reasons, it may be preferable to use numerical integration. In doing so, the width of the summation elements, $\Delta\xi = h$, must be sufficiently small to ensure that all oscillations of the phase term B are accurately sampled. On the other hand, the computations would be unnecessarily complicated if h is too small. A general rule for choosing h can be established as follows. As we are dealing with diffraction, the period of the oscillations does not depend on the specific functional form of B but only on the path length $P_0Q + QP$, where $P_0 = (x_0, z_0)$ is the source point shown in Fig. 1. When the point Q is moved along the x -axis by an increment h , this path length changes by

$$\begin{aligned} \delta &= \sqrt{z_0^2 + (\xi + h - x_0)^2} - \sqrt{z_0^2 + (\xi - x_0)^2} \\ &\quad - \sqrt{z^2 + (\xi + h - x)^2} + \sqrt{z^2 + (\xi - x)^2} \\ &\approx h \left(\frac{\xi - x_0}{|z_0|} + \frac{\xi - x}{z} \right), \end{aligned} \quad (2)$$

where it is assumed that $(\xi - x_0)^2 \ll z_0^2$, $(\xi - x)^2 \ll z^2$, and terms in h^2 are ignored. Hence it follows from the quarter-wave criterion that reliable results can be expected when h is chosen so that $\delta < \lambda/4$.

We now divide the aperture halfwidth a into N summation elements bounded by equidistant points Q_n , as illustrated in Figs. 2a, b, and 4, and define

$$OQ_0 = x = mh, \quad Q_0Q_n = \xi - x, \quad Q_{n-1}Q_n = h = \frac{a}{N}, \quad (3a)$$

where O is the coordinate origin, Q_0 is the projection of the point of observation P onto the x -axis, and m, n , and N are integers. Accordingly, Eq. (1) can be replaced by the quadrature formula

$$U(x, z) \approx U_m = \frac{a}{N} \sum_{n \neq 0} A_{mn} B_{mn}, \quad (3b)$$

where

$$A_{mn} = A[(n - \tfrac{1}{2})h, mh, z], \quad B_{mn} = B[(n - \tfrac{1}{2})h, mh, z] \quad (3c)$$

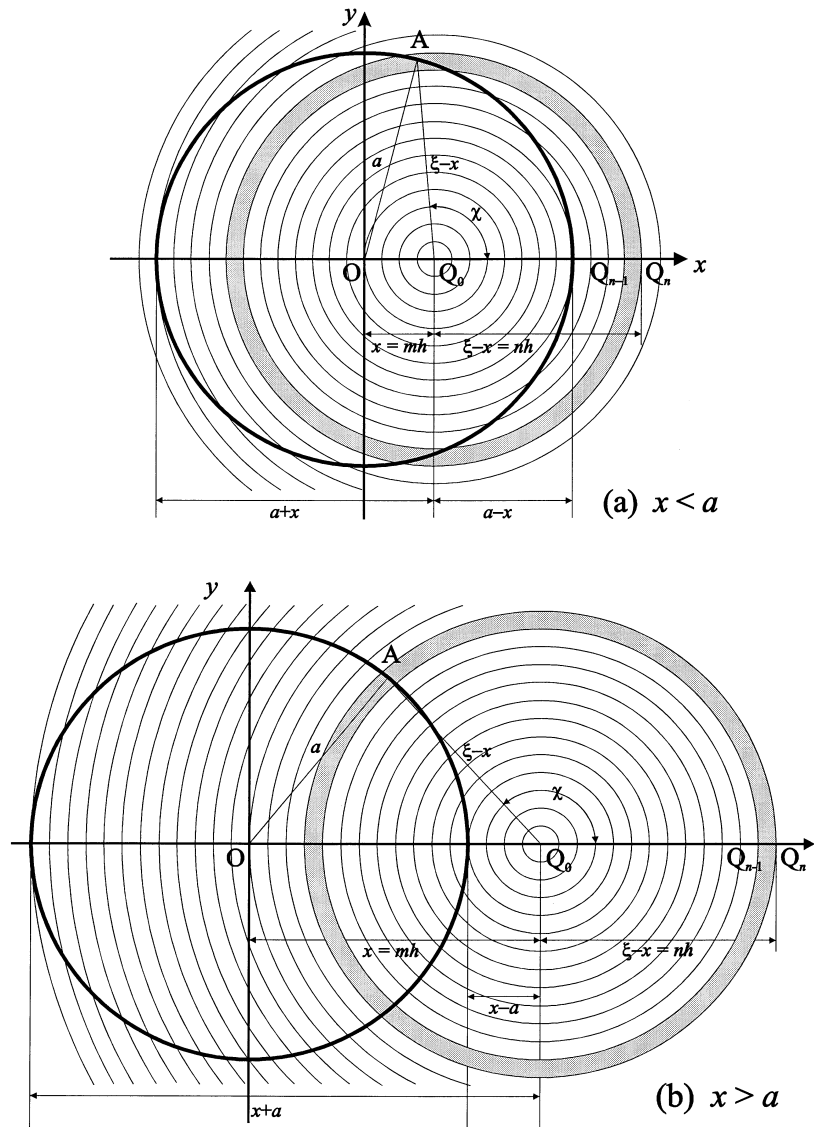


Fig. 2. Annular summation elements for a circular aperture (xy-plane). (a) Lit region. (b) Shadow region.

are the values of A and B at the midpoints of the summation elements and the limits of summation depend on the diffraction problem being considered. The value of N to be used in these formulae can be estimated from Eq. (2) by assuming a distant source ($|z_0| \gg z$) so that $\delta \approx h(\xi - x)/z = a^2 n / (N^2 z)$. If this is to be less than $\lambda/4$ for every summation element used for the computations, a good upper limit for n is $3N$,¹ and then one finds

$$N \geq \frac{12a^2}{\lambda z} = \frac{6u}{\pi} \approx 2u, \quad (3d)$$

where $u = ka^2/z$ is the familiar configuration parameter of Fresnel's diffraction theory for $|z_0| \gg z$. As it is well known that diffraction patterns pertaining to large values of u are highly structured [4], this result makes good sense in that it stipulates narrower summation elements when u increases. The accuracy of the numerical computations can also be improved by replacing the value of B_{mn} in Eq. (3c) with the value corresponding to Simpson's rule,

¹ The value pertains to a point of observation located one aperture halfwidth beyond the shadow boundary. A larger value should be used if necessary.

$$B_{mn} = \frac{1}{6}\{B[(n-1)h, mh, z] + 4B[(n-\frac{1}{2})h, mh, z] + B[nh, mh, z]\}. \quad (3e)$$

By means of trial computations, it was found that this substitution can result in a tenfold improvement of accuracy.

The above equations are intended for applications where closed solutions of Eq. (1) cannot be found and will be used in future research. In the remainder of the present paper, their validity will be demonstrated by numerical examples involving the Fresnel diffraction integral

$$U_F(P) = U_{\text{geom}}(P)\alpha_F(P), \quad \alpha_F(P) = \frac{-i}{\lambda z} \int dQ e^{ik(QP-z)}, \quad (4a)$$

where U_{geom} is the geometrical field in the absence of diffraction, α_F is the modification of the field by diffraction, and where

$$QP = \sqrt{(\xi-x)^2 + \eta^2 + z^2} \approx z + \frac{(\xi-x)^2 + \eta^2}{2z}. \quad (4b)$$

These expressions are valid in the paraxial Fresnel approximation for a distant source point P_0 and pertain to a point of observation $P = (x, 0, z)$ as in Fig. 1 whereas the point $Q = (\xi, \eta, 0)$ is assumed to lie anywhere in the aperture plane $z = 0$. In Secs. 2 and 3 of the paper, α_F will be reduced to a two-dimensional integral for the respective cases of circular apertures and apertures bounded by infinite straight lines (slit and half plane). The results of the numerical integration will be shown and compared to the corresponding exact solutions, which may be found in Ref. [5].

2. Circular Aperture

For a circular aperture of radius a , Fresnel's integral in Eq. (4a) can be reduced to a single integral by assuming annular area elements dQ which are centered on the projection Q_0 of the point of observation onto the aperture plane, as indicated in Figs. 2a and 2b. With O as the aperture center, $x = OQ_0 > 0$ and $\xi = OQ > x$ for points to the right of Q_0 , the distance QP defined by Eq. (4b) will then be constant and equal to

$$QP = z + (\xi-x)^2/2z \quad (5a)$$

everywhere on dQ and the integration can be carried out over ξ alone. As the annular area elements are eccentric to the aperture they are, in general, partially obstructed by the aperture rim so that their effective areas will be

given by

$$dQ = 2\pi(1 - \chi/\pi)d(\xi-x)(\xi-x), \quad (5b)$$

where $\chi = \angle AQ_0Q$ is the semi-angle subtended at Q_0 by the intersection of the area elements with the aperture rim, as indicated in the Figs. 2a and 2b.

When Q_0 lies inside the aperture ($x \leq a$, as in Fig. 2a), the innermost area elements with radii $\xi-x \leq a-x$ are unobstructed and fully contained in the aperture ($\chi=0$), and the outermost elements with radii $\xi-x > a+x$ are fully obstructed ($\chi=\pi$). For the intermediate elements the angle χ is found by applying the cosine theorem to the triangle OAQ_0 shown in the figure, so that

$$\cos\chi = \frac{a^2 - x^2 - (\xi-x)^2}{2x(\xi-x)}. \quad (5c)$$

Thus, upon substitution of Eqs. (5a) and (5b) into Eq. (4a) and noting that $2\pi/(\lambda z) = k/z = u/a^2$,

$$\begin{aligned} \alpha_F(x) &= \frac{-iu}{a^2} \left[\int_0^{a-x} d(\xi-x)(\xi-x)e^{ik(\xi-x)^2/(2z)} \right. \\ &\quad \left. + \int_{a-x}^{a+x} d(\xi-x)(\xi-x)(1 - \chi/\pi)e^{ik(\xi-x)^2/(2z)} \right] \\ &= [1 - e^{iu(a-x)^2/(2a^2)}] \\ &\quad - \frac{iu}{a^2} \int_{a-x}^{a+x} d(\xi-x)(\xi-x)(1 - \chi/\pi)e^{ik(\xi-x)^2/(2z)}, \quad x \leq a, \end{aligned} \quad (5d)$$

where the first integral was reduced to an elementary expression by substituting $ik(\xi-x)^2/2z$ as the integration variable. When Q_0 lies outside the aperture ($x \geq a$, as in Fig. 2b), the inner elements with radii $\xi-x \leq x-a$ and the outer elements with radii $\xi-x \leq x+a$ are all fully obstructed ($\chi=\pi$). In the intermediate region, Eq. (5c) applies once again² and we have

$$\begin{aligned} \alpha_F(x) &= \frac{-iu}{a^2} \int_{x-a}^{x+a} d(\xi-x)(\xi-x)(1 - \chi/\pi)e^{ik(\xi-x)^2/(2z)}, \\ &\quad x \geq a. \end{aligned} \quad (5e)$$

The integrals in Eqs. (5d) and (5e) can now be identified with Eq. (1), with

² Although χ is discontinuous on crossing the shadow boundary, Eqs. (5d) and (5e) are equivalent and give identical results for $x = a$.

$$A(\xi, x, z) = \frac{-iu(\xi - x)(1 - \chi/\pi)}{a^2},$$

$$B(\xi, x, z) = e^{ik(\xi-x)^2/(2z)}, \quad (6a)$$

so that, according to Eqs. (3a-c),

$$\alpha_F(P) \approx a_m = [1 - e^{iu(N-m)^2/(2N^2)}] - \frac{iu}{N^2} \sum_{n=N-m+1}^{N+m} n(1 - \chi_{mn}/\pi) e^{iu(n-\frac{1}{2})^2/(2N^2)}, \quad m \leq N \quad (6b)$$

$$\alpha_F(P) \approx \alpha_m = -\frac{iu}{N^2} \sum_{n=m-N}^{m+N} n(1 - \chi_{mn}/\pi) e^{iu(n-\frac{1}{2})^2/(2N^2)}, \quad (6c)$$

where

$$\chi_{mn} = \cos^{-1} \left[\frac{N^2 - m^2 - (n - \frac{1}{2})^2}{2m(n - \frac{1}{2})} \right]. \quad (6d)$$

The use of these equations on a personal computer is simple. As an example, Fig. 3 compares the numerical values of $|\alpha_m|^2$ obtained for $u = 5\pi$ and $N = 16$ to the exact results given by the Fresnel-Lommel theory. The agreement is good and improves when larger values of N are used, as indicated in the left-hand column of Table 1. The values listed in the table are the maximum errors encountered in the range $m \leq 3N$. In this particular case they occurred near the center of the profile ($m \leq 0.5N$).

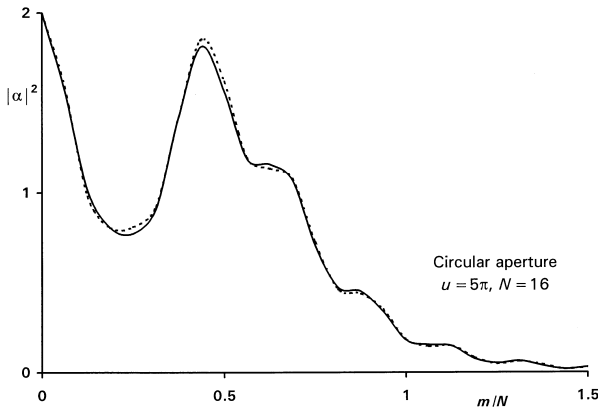


Fig. 3. Approximate (—) and exact (----) diffraction profiles of a circular aperture.

3. Apertures Bounded by Infinite Straight Lines

For a plane aperture of width $(l + r)$ which is bounded by infinite straight lines as in Fig. 4, the reduction of the integral of Eq. (4a) to two dimensions is

Table 1. The largest errors in diffraction profiles computed from Eqs. (6a), (6b), (6c), (10a), (10b), and (11b).

| N | Circular aperture ($u=5\pi$) | Slit ($u=5\pi$) | M | Half plane |
|-----|--------------------------------|----------------------|-----|----------------------|
| 16 | 6.0×10^{-2} | 2.3×10^{-2} | 8 | 1.9×10^{-2} |
| 32 | 2.3×10^{-2} | 5.6×10^{-3} | 16 | 4.3×10^{-3} |
| 64 | 6.4×10^{-3} | 1.4×10^{-3} | 32 | 1.1×10^{-3} |
| 128 | 2.5×10^{-3} | 3.6×10^{-4} | 64 | 2.7×10^{-4} |
| 256 | 9.6×10^{-4} | 9.0×10^{-5} | 128 | 6.6×10^{-5} |
| 512 | 3.5×10^{-4} | 2.3×10^{-5} | 256 | 1.7×10^{-5} |

readily achieved by choosing cartesian coordinates so that the y -axis is parallel to the aperture edges. Letting $P = (x, 0, z)$ as before, this leads to

$$\begin{aligned} \alpha_F(P) &= \frac{-i}{\lambda z} \int_{-\infty}^{\infty} d\eta e^{ik\eta^2/2z} \int_{-l}^r d\xi e^{ik(\xi-x)^2/(2z)} \\ &= \frac{-ie^{i\pi/4}}{\sqrt{\lambda z}} \int_{-(l+x)}^{(r-x)} d(\xi - x) e^{ik(\xi-x)^2/(2z)}, \end{aligned} \quad (7a)$$

where

$$\int_{-\infty}^{\infty} d\eta e^{ik\eta^2/(2z)} = \sqrt{2\lambda z} F(\infty) = \sqrt{\lambda z} e^{i\pi/4} \quad (7b)$$

and $F(\infty) = e^{i\pi/4}$ is the complex Fresnel integral at infinity.

The amplitude term $A(\xi, x, z)$ in Eq. (1) is now given by the factor $-ie^{i\pi/4}/\sqrt{\lambda z}$ that appears outside the integral of Eq. (7b) so that, on letting $l = Lh$ and $r = Rh$ as indicated in Fig. 4 and using Eqs. (3a) to (3d), we find

$$\alpha_m = \frac{-ihe^{i\pi/4}}{\sqrt{\lambda z}} \sum_{n=-(L+m)}^{R-m} e^{ik(n-\frac{1}{2})^2 h^2/(2z)}, \quad n \neq 0. \quad (8a)$$

It follows immediately that, if α_m is known and P is moved to the right or left by $\pm h$, the new value of α_m is given by the recursion formula

$$\alpha_{m\pm 1} = \alpha_m + \frac{ih e^{i\pi/4}}{\sqrt{\lambda z}} [e^{ik(L+m+\frac{1}{2})^2 h^2/(2z)} - e^{ik(R-m-\frac{1}{2})^2 h^2/(2z)}], \quad (8b)$$

which illustrates in a very instructive manner how the diffraction pattern changes when the point of observation is moved, so that one portion of the wavefront is uncovered by the aperture and another portion is covered. The recursion formula Eq. (8b) is convenient for practical computations as it allows the computation of successive values of α_m without performing the summation of Eq. (8a) for every point. In the examples given below, a starting value for α_m is obtained from known

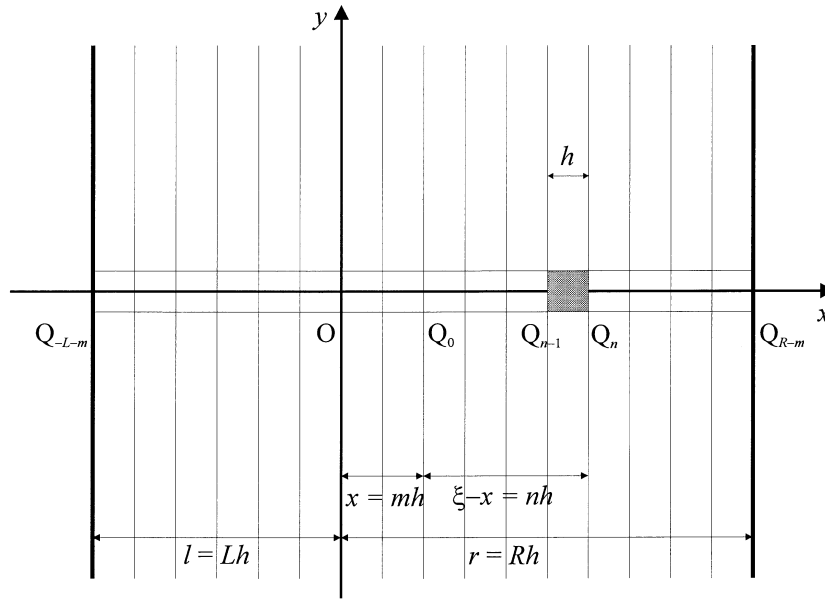


Fig. 4. Rectangular summation elements for apertures bounded by straight lines (xy-plane).

closed solutions, so that there is no need at all to perform a summation. This use of an exact starting value also improves the accuracy of the computations because it forces the initial error to be zero.

3.1 Slit

For a slit of width $2a$, we define $l = r = a$, $L = R = N$ and $h = a/N$, so that

$$\frac{h}{\sqrt{\lambda z}} = \frac{1}{N} \sqrt{\frac{u}{2\pi}}, \quad \frac{kh^2}{z} = \frac{u}{N^2}, \quad (9a)$$

$$\alpha_{m+1} = \alpha_m - \frac{i}{N} \sqrt{\frac{u}{2\pi}} e^{i\pi/4} [e^{iu(N+m+\frac{1}{2})^2/(2N^2)} - e^{iu(N-m-\frac{1}{2})^2/(2N^2)}], \quad (9b)$$

where it is noted that it is sufficient to perform the computations for $m > 0$ as the diffraction pattern is symmetrical.

On account of the trigonometric identity

$$\begin{aligned} e^{i\alpha} - e^{i\beta} &= e^{i(\alpha+\beta)/2} [e^{i(\alpha-\beta)/2} - e^{-i(\alpha-\beta)/2}] \\ &= 2ie^{i(\alpha+\beta)/2} \sin[(\alpha-\beta)/2], \end{aligned} \quad (10a)$$

this can be transformed into the following expressions, which are convenient for practical computations:

$$\alpha_{m+1} = \alpha_m + \frac{1}{N} \sqrt{\frac{2u}{\pi}} e^{iX_m} \sin Y_m, \quad (10b)$$

$$X_m = \frac{\pi}{4} + \frac{u}{2} \left[1 + \frac{(m + \frac{1}{2})^2}{N^2} \right], \quad Y_m = \frac{u(m + \frac{1}{2})}{N}. \quad (10c)$$

As an example, Fig. 5 shows the approximate and exact diffraction profiles of a slit for $u = 5\pi$, $N = 16$, and using the well-known Fresnel solution, $\alpha_F(0) = (1 - i)F(\sqrt{u/\pi})$, as the starting value. The two curves resemble each other closely, the largest errors being on the order of 0.023 and occurring near $m/N = 0.5$. The improvement of accuracy achieved by using larger values of N is shown in the center portion of Table 1.

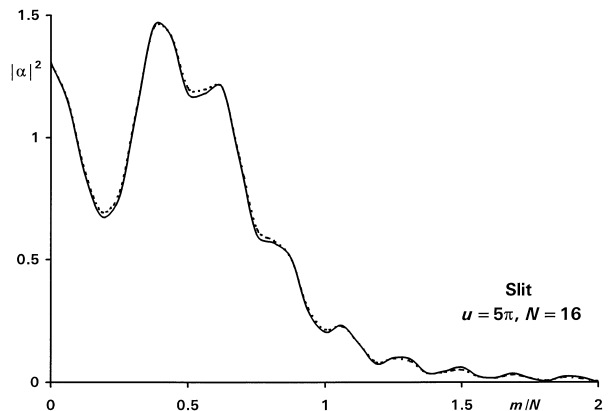


Fig. 5. Approximate (—) and exact (----) diffraction profiles of a slit.

3.2 Half Plane

On letting $L = 0$ and $R = \infty$, Eqs. (8a) and (8b) apply to a diffracting straight edge which coincides with the y -axis of Fig. 4. As there is no aperture edge on the right, the last term of Eq. (8b) is now absent and one obtains

$$\alpha_{m\pm 1} = \alpha_m \mp \frac{ih e^{i\pi/4}}{\sqrt{\lambda z}} e^{ik(m\pm 1 - \frac{1}{2})^2 h^2/(2z)}. \quad (11a)$$

The previous definition of h as a given fraction of aperture width is no longer applicable but can be replaced by taking h as a certain fraction, say $h = \sqrt{\lambda z}/M$, of the width of the first Fresnel zone. Hence it follows easily that

$$\alpha_{m\pm 1} = \alpha_m \mp \frac{i}{M} e^{iX_{m\pm 1}}, \quad X_{m\pm 1} = \pi \left[\frac{1}{4} + \frac{(m \pm 1 - \frac{1}{2})^2}{M^2} \right]. \quad (11b)$$

Figure 6 shows the diffraction pattern computed in this manner for $M = 4$, using the well-known Fresnel solution, $\alpha_F(0) = 1/2$, as the starting value. The agreement with the exact result is within ± 0.025 for $-4 \leq m/M < 1$, but considerably worse beyond these limits. These discrepancies are reduced when larger values of M are used, as may be seen from the right-hand side of Table 1.

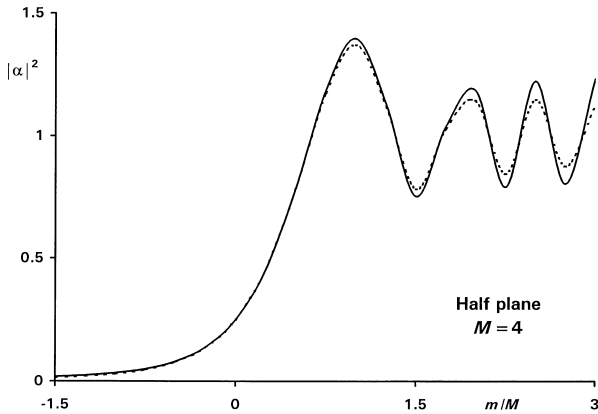


Fig. 6. Approximate (—) and exact (----) diffraction profiles of a half plane.

4. Appendix A. Evaluation of Eq. (7a) by the Method of Stationary Phase

Equation (7a) is of the form

$$\alpha_F(x) = \int_{-(l-x)}^{(r-x)} d\phi A e^{ikg(\phi)},$$

where $\phi = \xi - x$, $A = -ie^{i\pi/4}/\sqrt{\lambda z}$, $g(\phi) = \phi^2/z$. The only stationary point, defined by $g'(\phi) = \phi/z = 0$, occurs at $\phi_0 = 0$ so that, according to the Appendix of Ref. [2], the stationary-phase value of the integral is

$$\alpha_F(x) \approx \frac{\sqrt{\lambda} \varepsilon}{\sqrt{|g''(\phi_0)|}} A e^{ikg(\phi_0)},$$

where $g''(\phi_0) = 1/z$ and $\varepsilon = e^{+i\pi/4}$ according as $g''(\phi_0)$ is positive or negative. Since $1/z$ is positive, this gives $\alpha_F \approx 1$ so that the “diffraction pattern” of a slit or half plane would be identically equal to one everywhere in the plane of observation.

5. References

- [1] J. Focke, Ber. Sächs. Akad. Wiss. **101**, 3 (1954).
- [2] E. Wolf and E. W. Marchand, J. Opt. Soc. Am. **54**, 587 (1964).
- [3] A. Schuster, Phil. Mag. **31**, 77 (1891).
- [4] M. Born and E. Wolf, Principles of Optics, 4th Ed., Pergamon Press, Oxford (1970) p. 372.
- [5] K. D. Mielenz, J. Res. Natl. Inst. Stand. Technol. **103**, 497 (1998).

About the author: Klaus D. Mielenz is a physicist and retired Chief of the Radiometric Physics Division of the NIST Physics Laboratory. The National Institute of Standards and Technology is an agency of the Technology Administration, U.S. Department of Commerce.